(12) PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. AU 200016676 B2
(10) Patent No. 776910

(54) Title
Modifying protein immunogenicity

(51)6 International Patent Classification(s)
C07K 014/435 C07K 016/46
C07K 014/315 G01N 033/563

(21) Application No: 200016676 (22) Application Date: 1999.12.08

(87) WIPO No: WO00/34317

(30) Priority Data

(31) Number	(32) Date	(33) Country
9826925	1998.12.08	GB
9902139	1999.02.02	GB

(43) Publication Date : 2000.06.26
(43) Publication Journal Date : 2000.08.24
(44) Accepted Journal Date : 2004.09.23

(71) Applicant(s)
Biovation Limited

(72) Inventor(s)
Francis Joseph Carr; Fiona Suzanne Adair; Anita Anne Hamilton ; Graham Carter

(74) Agent/Attorney
F.B. Rice and Co.,139 Rathdowne Street,CARLTON VIC 3053

(56) Related Art
WO 1998/052976

100 10/30

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C07K 14/435, 14/315, 16/46, G01N 33/563 — A2

(11) International Publication Number: WO 00/34317

(43) International Publication Date: 15 June 2000 (15.06.00)

(21) International Application Number: PCT/GB99/04119

(22) International Filing Date: 8 December 1999 (08.12.99)

(30) Priority Data:
9826925.1 8 December 1998 (08.12.98) GB
9902139.6 2 February 1999 (02.02.99) GB

(71) Applicant (*for all designated States except US*): BIOVATION LIMITED [GB/GB]; Investment House, 6 Union Row, Aberdeen AB10 1DQ (GB).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): CARR, Francis, Joseph [GB/GB]; Biovation Limited, Crombie Lodge, Aberdeen Science Park, Balgownie Drive, Aberdeen AB22 8GU (GB). ADAIR, Fiona, Suzanne [GB/GB]; Biovation Limited, Crombie Lodge, Aberdeen Science Park, Balgownie Drive, Aberdeen AB22 8GU (GB). HAMILTON, Anita, Anne [GB/GB]; Biovation Limited, Crombie Lodge, Aberdeen Science Park, Balgownie Drive, Aberdeen AB22 8GU (GB). CARTER, Graham [GB/GB]; Biovation Limited, Crombie Lodge, Aberdeen Science Park, Balgownie Drive, Aberdeen AB22 8GU (GB).

(74) Agents: SHEARD, Andrew, Gregory et al.; Kilburn & Strode, 20 Red Lion Street, London WC1R 4PJ (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MODIFYING PROTEIN IMMUNOGENICITY

(57) Abstract

It is known that proteins, or parts of them, may be rendered non- or less immunogenic to humans or other species by identifying one or more potential T cell epitopes and eliminating them by amino acid modification. Conventionally, certain epitopes may be retained in a protein sequence if the peptides constituting such epitopes are present in endogenous human protein, since they would be recognised as "self". However, it has now been found that even self epitopes may give rise to immune reactions. The invention provides for their elimination, for example by recombinant DNA technology, to render them more useful for administration to humans, for example for therapeutic or diagnostic purposes.

MODIFYING PROTEIN IMMUNOGENICITY

THE PRESENT INVENTION relates to proteins to be administered especially to humans particularly for therapeutic use but also for use in diagnostic tests. The invention particularly provides for proteins which are modified to be less immunogenic than the unmodified counterpart when used *in vivo*.

The invention particularly addresses the clinical need for a process by which the natural tendency of the host to mount an immune response against an administered protein is substantially reduced or eliminated. There are several examples where the administration of protein molecules offers therapeutic benefit. However the benefit is greatly reduced particularly where multiple doses of the therapeutic protein are required as the recipient immune system recognises and accelerates the elimination of the incoming therapeutic protein.

There are a number of therapeutic proteins whose therapeutic use is curtailed on account of their immunogenicity in man. For example, when murine antibodies are administered to patients who are not immunosuppressed, a majority of patients mount an immune reaction to the foreign material by making human anti-murine antibodies (HAMA). There are two serious consequences. First, the patient's anti-murine antibody may bind and clear the therapeutic antibody or immunoconjugate before it has a chance to bind to the tumour and perform its function. Second, the patient may develop an allergic sensitivity to the murine antibody and be at risk of anaphylactic shock upon any future exposure to murine immunoglobulin.

Several techniques have been employed to address the HAMA problem and thus enable the use in humans of therapeutic monoclonal antibodies (see, for example WO-A-8909622, EP-A-0239400, EP-A-0438310, WO-A-9109967). A common aspect of these methodologies has been the introduction into the therapeutic antibody, which in

general is of rodent origin, of significant tracts of sequence identical to that present in human antibody proteins. Such alterations are also usually coupled to alteration of particular single amino acid residues at positions considered critical to maintaining the antibody-antigen binding interaction. For antibodies, this process is possible owing to the very high degree of structural (and functional) conservatism between antibody molecules of different species. However for potentially therapeutic proteins where no structural homologue may exist in the host species (e.g. human) for the therapeutic protein, such processes are not applicable. Furthermore, these methods have assumed that the general introduction of human sequence will render the re-modelled antibody non-immunogenic. It is known however, that certain short peptide sequences ("T cell epitopes") can be released during the degradation of proteins within cells and subsequently presented by molecules of the major histocompatability complex (MHC) in order to trigger the activation of T cells. For peptides presented by MHC class II, such activation of T cells can then give rise to an antibody response by direct stimulation of B cells to produce such antibodies. None of the previous methods have addressed the elimination or avoidance of such epitopes in the final therapeutic molecule as a means to reduce or eliminate the antibody response against proteins. Nor have previous methods considered the elimination of peptides presented by MHC class I which can trigger a cytotoxic T cell response leading to cell killing. Such killing of cells causes the release of cellular components including proteins which can, in turn, activate specialist cells which are highly active in protein processing and MHC presentation and also in release of inflammatory cytokines as a result of such activation. As a result, an inflammatory environment can be created which promotes the more active uptake and processing of the protein for therapeutic use thus facilitating induction of an antibody response against the protein.

The elimination of T cell epitopes from proteins has been previously disclosed (see WO-A-9852976) whereby such potential T cell epitopes are defined as any peptide within the protein sequence with the ability to bind to MHC class II molecules. Such potential epitopes are measured by any computational or physical method to establish

MHC binding. Implicit in the term "T cell epitope" is an epitope which is recognised by the T cell receptor and which can at least in principle cause the activation of these T cells. It is however usually understood that certain peptides which are found to bind to MHC class II molecules may be retained in a protein sequence because such peptides are recognised as "self" within the organism into which the final protein is administered.

In practice, soluble proteins introduced into autologous organisms do sometimes trigger an immune response resulting in development of host antibodies which bind to the soluble protein. One example is interferon alpha 2 to which a percentage of human patients make antibodies despite the fact that this protein is produced endogenously.

The present invention is based on the discovery by the inventors that MHC-binding peptides within autologous proteins can trigger immune responses in live organisms to those proteins even when the specific protein is endogenously produced. A possible explanation for this is that the doses of such proteins administered are much higher than normal providing MHC-peptide formation which can activate T cells. Alternatively, the physiological environment into which such administered proteins are presented is one favourable to efficient antigen processing and peptide-MHC formation at higher levels than usually encountered, for example in inflammatory situations.

According to a first aspect of the invention, there is provided a method of rendering a protein, or part of a protein, non-immunogenic, or less immunogenic, to a given species, the method comprising:

(a) determining at least part of the amino acid sequence of the protein;

(b) identifying in the amino acid sequence one or more potential epitopes for T

cells ("T cell epitopes") which are found in an endogenous protein of the given species; and

(c) modifying the amino acid sequence to eliminate at least one of the T cell epitopes identified in step (b) thereby to reduce the immunogenicity of the protein or part thereof when exposed to the immune system of the given species.

The given species will usually be human. In step (c), one or more actual epitopes or potential epitopes may be eliminated.

Therefore, the present invention provides a modified method for creating proteins which trigger a reduced or absent immune response whereby one or more MHC binding peptides which are also found in the autologous organism's endogenous proteins are modified to reduce or eliminate binding to MHC molecules. Thus, in practice, no assumption is made about which peptides might be tolerated by the organism (unless data is available which indicates that tolerance is in place). Overall, the invention encompasses the creation of protein molecules with reduced or absent immunogenicity in live organisms where one or more peptides which bind to MHC molecules are eliminated from the protein molecule. As such, the invention is an improvement on the prior art which focuses only on the removal of potential T cell epitopes from proteins with the assumption that the organism is tolerant to self proteins or peptide sequences. In particular, the invention includes a range of human and non-human proteins which are modified by modification of one or more MHC binding peptides within the sequences. Such MHC binding peptides can include peptides which bind to MHC class II molecules and/or peptides which bind to MHC class I.

The aspect of the invention described above is preferred to the alternative strategy of separately immunising the organism with such peptides in order to induce tolerance or

the alternative of administering analogues or fragments of the protein which resist binding or uptake into cells (especially antigen-presenting cells), for example by elimination of the cell binding site in the protein. Tolerising the organism to the peptide in this way (*e.g.* the use of MHC binding peptides identified from a syngeneic protein which can be used to induce tolerance in the host prior to administration of the entire protein) would require two therapeutic molecules, one to perform the main protein function and the other to induce the tolerance, which would not be favoured from the regulatory point of view.

Any method of identifying – in which term is included predicting – one or more potential T cell epitopes can be used in the invention. Acceptable methods may be computational or physical and include measuring of the binding of MHC-peptide complexes to T cell receptors, the testing of potential T cell epitopes in transgenic mice expressing human MHC molecules, the testing of potential T cell epitopes in mice reconstituted with human antigen-presenting cells and T cells in place of their endogenous cells, and the testing of potential T cell epitopes for stimulation of T cells *in vitro* via presentation of peptides on MHC molecules using syngeneic antigen-presenting cells.

As well as identifying in the amino acid sequence one or more potential T cell epitopes which are found in an endogenous protein of the given species, a method of the invention may, and typically will, involve additionally identifying and eliminating one or more potential T cell epitopes which are *not* found in an endogenous protein of the given species, since such epitopes are even more likely to be immunogenic.

The present invention is based on a new concept in development of therapeutic proteins. It has been recognised previously that, with exogenous T cell dependent proteins, helper T cell epitopes are important to development of a significant immune response to the autologous protein. Such T cell epitopes function through the internal

processing of the protein releasing peptides which complex with MHC class II molecules and which are then presented on the surface of appropriate cells for potential binding to receptors on T cells. However, it is very difficult to predict which peptides within the protein will be processed appropriately such that MHC binding can occur and it is also a factor that many processed peptides will not bind to all allotypic variants of MHC class II molecules (MHC restriction). Furthermore, in any given living organism, it is difficult to predict whether a T cell response will actually be triggered by a given peptide presented on MHC class II because T cells may have been tolerised to such epitopes or may not have the repertoire of T cell receptors to bind to the MHC class II-peptide complex. Due to these complicating factors, it has previously been impractical to develop proteins with reduced or absent immunogenicity because of the difficulty in predicting actual T cell epitopes. The invention herein primarily takes the new approach of creating improved therapeutic proteins by removal of potential rather than actual T cell epitopes (usually defined by binding to MHC molecules) such that certain peptides within the molecule which are not actually immunogenic may be altered in addition to those immunogenic peptides. For a therapeutic molecule, preferably all of the potential T cell epitopes are removed whilst retaining activity of the protein. Preferably, this involves judicious choice of the amino acid substitutions enabling removal of the T cell epitopes and usually will involve the testing of a range of variant molecules with different amino acid substitutions.

Not all identified potential T cell epitopes which are found in an endogenous protein need be eliminated. For example, immune responses in general are not mounted to autologous circulating proteins, such as immunoglobulins and other serum proteins such as serum albumin. So potential T cell epitopes which are found for example in germline immunoglobulin variable region protein sequences of the given species (which will generally be human) may sometimes be ignored. However, an epitope considered as not generally available to the immune system to gain tolerance may be identified as a potential epitope for elimination according to the method of the present

invention. Examples of such include epitopes include those of intracellular proteins, such as nuclear proteins and integral membrane proteins.

The invention therefore provides proteins which have been altered to reduce their immunogenicity as well as a general method of altering such proteins to reduce their immunogenicity. A major principle of the present invention is that proteins are altered by identification of potential T cell epitopes and their subsequent alteration within the proteins in order to eliminate such potential epitopes. Optionally, epitopes recognised by host antibodies ("B cell epitopes") can also be removed if these can be identified for example with the aid of patient antisera or if the surface residues in a non-autologous protein can be altered to those of a related autologous protein endogenous to the host organism without losing all of the desired activity of the non-autologous molecule.

After elimination of one or more T cell epitopes, the protein may be tested for desired activity. It may retain its full activity, or it may retain a sufficient proportion of its original activity to be useful. In some circumstances, the activity may be altered either beneficially or at least in an acceptable way. Proteins entirely lacking in useful function post modification may be discarded.

Proteins of the present invention include those which have potential clinical use in humans or veterinary use in animals, especially where such uses involve multiple administrations of the proteins. They may be non-autologous or autologous. Proteins of the present invention include proteins with an enzymatic activity which has a beneficial therapeutic effect, proteins which are used to convert inactive drugs to active drugs within living organisms, proteins which are used to vaccinate whereby certain immunogenic epitopes are undesirable, proteins which perform as carriers of other molecules within the living organism and proteins which bind to other molecules within or introduced within the living organism in order to alter the biodistribution of the other molecules. Amongst the examples of non-autologous proteins with potential

benefit for *in vivo* human use are the following: Thrombolytic proteins streptokinase and staphylokinase; pro-drug activating enzymes such as asparaginase and carboxypeptidase G2; toxins of plant, fungal and bacterial origin such as ricin, saponin, pokeweed antiviral protein, mistletoe lectin, cholera toxin, pertussis toxin, diphtheria toxin; bacterial phospholipase C (e.g. PE33, PE35, PE37, PE38, PE40); ribotoxins such as alpha-sarcin, clavin, mitogillin, restrictocin and bryodin-1 and 2. Other significant non-autologous molecules of proven and potential therapeutic benefit include streptavidin, cobra venom factor, insulin, collagen, non-autologous antibodies, non-autologous MHC molecules and non-autologous T cell receptor molecules. Autologous proteins of proven and potential therapeutic value include IL-2, interferon α and β, granulocyte macrophage colony stimulating factor (GMCSF), granulocyte colony stimulating factor (GCSF), tissue plasminogen activator (t-PA), insulin, Factor VIII, Factor IX, erythropoietin (EPO), pituitary growth hormone and megakaryocyte growth and development factor (MGDF) and derivatives of these proteins produced by recombinant methods. Other examples of proteins include recombinant molecules derived or comprised of single or multiple domains from autologous or non-autologous proteins, recombinant proteins engineered to change or modify function or protein sequences derived from any of the above molecules.

A typical protocol within the general method of the present invention comprises the following steps:

I. Determining the amino acid sequence of the protein or a part thereof (if modification only of a part is required);

II. Identifying potential T cell epitopes within the amino acid sequence of the protein by any method including determination of the binding of peptides to MHC molecules, determination of the binding of peptide:MHC complexes to the T cell receptors from the species to receive the therapeutic protein, testing of the protein or peptide parts thereof using transgenic animals with the MHC

molecules of the species to receive the therapeutic protein, or testing with transgenic animals reconstituted with immune system cells from the species to receive the therapeutic protein;

III. By genetic engineering or other methods for producing modified proteins, altering the protein to remove one or more of the potential T cell epitopes and producing such an altered protein for testing;

IV. (optionally) Within step III., altering the protein to remove one or more of the potential B cell epitopes;

V. Testing altered proteins with one or more potential T cell epitopes (and optionally B cell epitopes) removed in order to identify a modified protein which has retained all or part of its desired activity but which has lost one or more T cell epitopes.

Potential T-cell epitopes herein are defined as specific peptide sequences which either bind with reasonable efficiency to MHC class II molecules (or their equivalent in a non-human species), or which in the form of peptide:MHC complexes bind strongly to the T cell receptors from the species to receive the therapeutic protein or which, from previous or other studies, show the ability to stimulate T-cells via presentation on MHC class II. The method of the present invention recognises that an effective T cell-dependant immune response to a foreign protein requires activation of the cellular arm of the immune system. Such a response requires the uptake of the therapeutic (foreign) protein by antigen presenting cells (APCs). Once inside such cells, the protein is processed and fragments of the protein form a complex with MHC class II molecules and are presented at the cell surface. Should such a complex be recognised by binding of the T cell receptor from T-cells, such cells can be, under certain conditions, activated to produce stimulatory cytokines. The cytokines will elicit differentiation of B-cells to full antibody producing cells. In addition, such T cell

responses may also mediate other deleterious effects on the patient such as inflammation and possible allergic reaction.

It is understood that not all peptide sequences will be delivered into the correct MHC class II cellular compartment for MHC class II binding or will be suitably released form a larger cellular protein for subsequent MHC class II binding. It is further understood that even such peptides which are presented by MHC class II on the surface of APCs may not elicit a T-cell response for reasons including a lack of the appropriate T-cell specificity, tolerance by the immune system to the particular peptide sequence or the low affinity of the MHC-peptide complex for T cell receptor.

The present invention provides for removal of human (or other given species) potential T-cell epitopes from the therapeutic protein whereby the primary sequence of the therapeutic protein can be analysed for the presence of MHC class II binding motifs by any suitable means. For example, a comparison may be made with databases of MHC-binding motifs such as, by searching the "motifs" database at the world-wide web site http://wehih.wehi.edu.au/mhcpep/ (formerly wehil.wehi.edu.au or wehi.wehi.edu.au). Alternatively, MHC class II binding peptides may be identified using computational threading methods such as those devised by Altuvia *et al.* ((*J. Mol. Biol.* 249 244-250 (1995))). Similar methods may be used to identify and remove epitopes which bind to MHC class I.

Having identified potential given species (*e.g.* human) T-cell epitopes, these epitopes are then eliminated by alteration of one or more amino acids, as required to eliminate the T-cell epitope. Usually this will involve alteration of one or more amino acids within the T-cell epitope itself. This could involve altering an amino acid adjacent to the epitope in terms of the primary structure of the protein or one which is not adjacent in the primary structure but is adjacent in the secondary structure of the molecule. The usual alteration contemplated will be amino acid substitution, but in certain circumstances amino acid deletion or addition will be appropriate. In some instances,

selection of the appropriate amino acid substitution at any particular point in the therapeutic protein primary structure, can be made with reference to an homologous protein primary structure. This is particularly the case where a repertoire of homologous genes exist, as for example with immunoglobulins, or a single homologous gene for the therapeutic protein exists, as is the case for the cobra venom factor protein and human complement factor C3 detailed in Example 1. However in other cases, there may not be an homologous protein to the therapeutic protein as is the case for potential therapeutics such as the thrombolytic agent staphylokinase. In such cases amino acids may be selected on the basis of similar size and/or charge, or more preferably and where structural data for the protein exist in combination with or with reference to *in silico* protein modelling techniques. In some instances amino acid substitutions may be made to prevent proteolytic cleavage of a protein and thereby inhibit binding of a peptide to MHC class II. Examples of protease cleavage sites include those reported by van Noort & van der Drift (*J. Biol. Chem.* 264 14159-14164 (1989)) and van Noort *et al* (*Eur. J. Immunol.* 21 1989-1996 (1991)). Examples of amino acids which prevent protease digestion have been identified e.g. Kropshofer *et al.*, *J. Immunol.* 151 4732-4742 (1993).

In the method of the present invention, usually a number of variants of altered proteins will be produced and tested for retention of the desired activity. Where it is desirable to maximise the removal of potential T cell epitopes from the protein, it will be common to create a range of variants including some with potential T cell epitopes remaining in the molecule. It will be recognised that with certain protein molecules, it is difficult to radically alter the molecule and to retain full activity and so judicious use of molecular modelling of variants and the testing of a maximum number of variants is desirable. For molecular modelling, standard commercially available software packages can be used to model the protein structure using as starting point either of a crystal structure or a model built from homology or protein folding prediction. Information on parts of the protein involved in imparting the molecule with its activity will assist in modelling variants enabling best choice of altered amino acids which

remove a potential T cell epitope (or optional B cell epitope) whilst retaining activity.

In the method of the present invention where a number of variants of altered proteins will be produced and tested for retention of the desired activity, it is desirable to have a high-throughput method for screening large numbers of variants. Such methods include *in vivo* methods for expression of gene variants in cells such as bacteria with rapid isolation of the proteins for example using an antibody to a conserved region of the proteins, or via a protein purification tag included in the protein sequence. As an alternative, *in vitro* methods such as *in vitro* transcription and translation can be used in some cases. Where the protein binds to another molecule, display methods such as phage display and ribosome display can be used in order to select out variants which retain binding activity.

In the practice of the present invention, specific amino acid alterations can be conducted by recombinant DNA technology, so that the final molecule may be prepared by expression from a recombinant host using established methods. However, the use of protein chemistry or any other means of molecular alteration is not excluded in the practice of the invention. Whilst the present invention provides a method for removal of potential T cell epitopes (and optional B cell epitopes) from protein molecules, the method does not exclude testing of variant molecules produced within the invention for actual T cell epitope activity including testing of the protein or peptide parts thereof using transgenic animals with the MHC molecules of the species to receive the therapeutic protein, or testing with transgenic animals reconstituted with immune system cells from the species to receive the therapeutic protein.

Whilst such methods are not yet routine, *in vitro* T cell assays can be undertaken whereby the protein can be processed and presented on MHC class II by appropriate antigen-presenting cells (APCs) to syngeneic T cells. T cell responses can be measured by simple proliferation measurements (especially if the APCs have been irradiated or otherwise treated to prevent proliferation) or by measuring specific

cytokine release. In order to account for different MHC class II allotypes, a range of *in vivo* assays will usually be required in order to broadly test for T-cell epitopes. Alternatively, transgenic animals equipped with human (or the desired species) MHC class II molecules could be used to test for T-cell epitopes especially where the host MHC class II repertoire has been removed and especially where one or more other host accessory molecules in APC/T-cell interaction have also been replaced with human (or the desired species) such as CD4 on T-cells.

In a further aspect, the invention provides molecules resulting from method as described above. Such molecules may be useful for treating or preventing a disease or condition. The invention also extends to the use of such molecules in *in vivo* and *in vitro* diagnosis and for human or veterinary use. Preferred features of each aspect of the invention are as for each other aspect, *mutatis mutandis*.

The present invention provides a method of rendering an endogenous intracellular protein, or part of a protein, non-immunogenic, or less immunogenic, to a given species, the method comprising:

(a) determining at least part of the amino acid sequence of said endogenous intracellular protein;

(b) identifying in the amino acid sequence one or more potential epitopes for T cells ("T cell epitopes") which are found in said endogenous intracellular protein; and

(c) modifying the amino acid sequence to eliminate at least one of the T cell epitopes identified in step (b) thereby to reduce the immunogenicity of the protein or part thereof when exposed to the immune system of the given species.

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed before the priority date of each claim of this application.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

The invention is illustrated but not limited by the following examples. The examples refer to the following figures:

FIGURE 1 shows the protein sequence of mature cobra venom factor;

FIGURE 2 shows the protein sequence of an altered cobra venom factor-chain;

FIGURE 3 shows the protein sequence of an altered cobra venom factor-chain;

FIGURE 4 shows the protein sequence of an altered cobra venom factor-chain;

FIGURE 5 shows the protein sequence of streptokinase from *Streptococcus equisimilis;*

FIGURE 6 shows the protein sequence of an altered streptokinase molecule;

FIGURE 7 shows the protein sequence of mature staphylokinase;

FIGURE 8 shows the protein sequence of an altered staphylokinase molecule;

FIGURE 9 shows the protein sequence of wild-type bryodin 1;

FIGURE 10 shows the protein sequence of an altered bryodin 1 molecule;

FIGURE 11 shows the protein sequence of mature human interferon 2 ; and

FIGURE 12 shows the protein sequence of an altered human interferon 2 molecule.

Example 1

In the present example the cobra venom factor (CVF) protein is analysed for the presence of potential MHC binding motifs and a method disclosed for the removal of a number of these from the molecule.

CVF is the non-toxic complement-activating glycoprotein in cobra venom. The mature CVF protein consists of three polypeptide chains α, β and γ, produced by post-translational processing of a pre-pro protein of approximate molecular weight 150kDa (Fritzinger D.C. *et al.*, *Proc. Nat'l. Acad. Sci. USA* 91 12775-12779 (1994)). The α and β chains are both linked to the γ chain via single disulphide bridges. In addition, both the α chain and the γ chain have single intra chain disulphide bonds whilst the β chain has a complex arrangement of six further internal disulphide bonds. The mature protein has four potential sites for N-glycosylation of which only three are used

(Vogel, C.W & Muller-Eberhard *J. Immunol.* 18 125-133 (1984)).

In the preferred embodiment of the present invention, production of the reduced immunogenicity or altered version of the therapeutic protein is by recombinant DNA technology. For the present example, it is understood that the expedient of recombinant production of the mature CVF protein alone is likely to provide a considerable reduction in immunogenicity of the protein. This assertion stems from the knowledge that a degree of the immunogenicity arising from administration of CVF purified from *Naja naja* venom, arises from antibody responses to the particular pattern of glycosylation not encountered in mammalian proteins (Gowda, D.C *et al.*, *J. Immunol.* 152 2977-2986 (1994)). It is also understood that deglycosylated and also sialylated CVF that would arise from recombinant production of CVF, for example in a mammalian producer cell system, have equal functional activity with native mature CVF (Gowda, D.C *et al.* (1994) *ibid.*).

In humans CVF triggers the alternative (properdin) pathway of complement activation. CVF forms a bimolecular complex with factor B which is cleaved by factor D into Bb which remains bound to CVF to form the C3/C5 convertase CVF,Bb. This complex is functionally and structurally analogous to the human C3/C5 convertase formed during the activation of the alternative pathway. The clinical importance of CVF lies in the high relative stability of the CVF,Bb complex versus the native C3/C5convertase ($t_{1/2}$ of 7 hours versus $t_{1/2}$ of 90 seconds) and in particular, the inability of CVF,Bb to be disrupted and hence down regulated by factors H and I (Gowda, D.C *et al.*, (1994) *ibid.*). The net result of CVF administration therefore is to lead to the depletion of complement in human serum by continuous efficient and uncontrolled complement activation.

This property has been exploited as a research tool for *in vivo* and *in vitro* de-complementation (Cochrane C.G *et al.*, *J. Immunol.* 105 55 (1970)) and has also been exploited for the selective elimination of tumour cells as a conjugate with a tumour

targeting monoclonal antibody (Vogel, C.-W. & Muller-Eberhard H.J. *Proc. Nat'l. Acad. Sci. USA* 78 7707 (1981)). Availability of a non-immunogenic CVF would have considerable importance as a potential therapeutic agent for the depletion of complement in the plasma of patients undergoing organ xenotransplantation or in the construction of second generation antibody conjugates for targeted activation of complement in cancer patients.

Method for identification of potential MHC class II binding motifs in cobra Venom Factor:

The sequence of CVF was identified from the GenBank database. The sequence with accession number U09969 was used throughout. The protein sequences of the mature α, β and γ chains were identified (Figure 1) and analysed individually for the presence of potential MHC class II binding motifs. Analysis was conducted by computer aided comparison to a database of MHC-binding motifs as resident on world wide web site http://wehih.wehi.edu.au/mhcpep/.

Results of the "searching" process on the α chain, indicate the presence of 638 potential MHC class II binding motifs. Of these, 525 matched sequences identified in a database of human germline immunoglobulin variable region protein sequences. These epitopes were not considered further on the basis that immune responses in general are not mounted to autologous circulating proteins such as immunoglobulins. This implies immunological tolerance to the potential T-cell epitopes present in the structure of the immunoglobulins (and indeed the majority of human proteins). Epitopes presented by non-autologous proteins such as CVF which are identical or similar to motifs present in immunoglobulin proteins are likely also to be tolerated and in practice may be retained through the de-immunisation process.

Following subtraction of the human immunoglobulin protein germline motifs, the remaining 113 potential epitopes in the α chain, were analysed individually for

similarity to non-immunoglobulin protein sequences. In practice, predicted anchor residues for each potential epitope were used in a consensus sequence search of human expressed proteins. The SwissProt and GenBank translated sequence databases were interrogated using commercially available software (DNAstar Madison, WI, USA). Epitopes identified in known circulating human proteins were not considered further and were therefore allowed to remain unchanged within the CVF α chain. An example of one such rejected potential epitope is given by the sequence FKPGMPY at positions 332-338 in the CVF α chain (numbering from first amino acid of the mature α chain). This sequence represents a predicted consensus binding motif for HLA-DRQB1*0301 with anchor residues underlined. Database searching using the consensus sequence FxxxMxY identifies only a single entry in a human protein sub-set of the SwissProt database corresponding to human serum biotinidase (SwissProt accession # A54362). An example of an epitope where no match to a human protein considered to be in the general circulation was found is provided by sequence YYQVGNNEL at position 501-509 in the CVF α chain. This sequence represents a potential epitope for presentation by HLA-DRB1*0101. Consensus sequence searching identifies only four human proteins containing this motif, of which three are nuclear proteins of differentiated tissues such as brain, and the fourth is an integral membrane protein. These may be considered as not generally available to the immune system to gain tolerance and therefore identify this as a potential epitope for elimination according to the method of the present invention. Similarly, a further potential HLA-DR1 binding motif was identified in the α chain. Peptide sequence YVVVQVTGP at positions 81-89. This motif identifies a single human nuclear protein DNA polymerase alpha (GenBank accession # M64481), in the same data set and was also therefore identified for modification by the method of the present invention. Following these processes, a total of 15 potential epitopes in the α chain were considered for removal by amino acid substitution.

Results of the "searching" process on the β chain, indicate the presence of 366

potential MHC class II binding motifs. Of these, 281 matched sequences identified in a database of human germline immunoglobulin variable region protein sequences and were not considered further. Following a second round of subtractions using a database of human serum protein sequences (as above), 18 potential epitopes were considered for removal.

Results of the "searching" process on the γ chain, indicate the presence of 267 potential MHC class II binding motifs. Of these, 219 matched sequences identified in a database of human germline immunoglobulin variable region protein sequences and were not considered further. Following a second round of subtractions using a data base of human serum protein sequences (as above), 9 potential epitopes were considered for removal.

The net result of these processes was to identify those residues within the CVF molecule which should be altered to eliminate potential MHC class II binding motifs. Individual amino acids within the predicted binding motifs were selected for alteration. With the object of maximising the likelihood of maintaining protein functional activity, in all cases conservative amino acid substitutions were chosen at any given site. In many cases, individual amino acid or even short strings (<5 consecutive residues) were substituted by reference to the published sequence of human complement factor 3 (C3) (GenBank accession # K02765) which shows areas of strong homology, including regions of identity, with CVF. In two instances, potential epitopes were eliminated from the β chain by insertion of additional amino acids. Amino acids were selected with reference to the human C3 protein.

Altered CVF α chain, β chain and γ chain sequences were compiled (Figures 2 to 4) and further analysed by database comparison, as previously, for confirmation of successful elimination of potential MHC class II binding motifs.

Method for construction of altered CVF molecules:

PCR primers CVF1for; 5'ATAAGAATGCGGCCGCATGGAGAGGATGGCTCTCT and CVF1rev; 5'ATAAGAATGCGGCCGCTATCATTGATTCTTCTGAAC were used to amplify a 4985bp fragment from a λgt11 cobra venom gland cDNA library (Fritzinger DC *et al., ibid.*). The PCR was conducted on a total library DNA preparation using high fidelity polymerase mix optimised for long distance PCR (*Advantage PCR Kit,* Clontech, Basingstoke, UK) and conditions suggested by the supplier. The PCR product was cloned into pcDNA3.1 (Invitrogen, Leek, The Netherlands) as a *Not*I restriction fragment using standard techniques (Sambrook J., Fritsch E.F. & Maniatis T. (eds.) in: Molecular Cloning a Laboratory Manual, Cold Spring Harbor Laboratory Press, NY, USA (1989)). The gene sequence was confirmed to be identical to database entries using commercially available reagent systems and instructions provided by the supplier (Amersham, Little Chalfont, UK). Site directed mutagenesis was conducted using synthetic oligonucleotides and the "quick-change" procedure and reagents from Stratagene (Cambridge, UK). Mutated (de-immunised) versions of the gene were confirmed by sequencing. Mutated CVF genes were transfected into CHO cells by electroporation. Stable transfectants were selected using G418, and clones secreting active CVF were selected using a C consumption assay (Gowda D.C *et al.* (1994) *ibid.*; Cochrane C.G. et al (1970) *ibid.*) Expressing clones were expanded and recombinant protein was purified from the culture supernatant using sequential column chromatography and methods as described (Vogel C-W *et al., J. Immunological Methods* 73 203-220 (1984)).

Example 2

The present invention details a process whereby potentially immunogenic epitopes within a non-autologous protein may be identified and offers methodology whereby such epitopes may be eliminated. It is understood that there are a number of proven therapeutic proteins for which their therapeutic use is curtailed on account of their immunogenicity in man. In the present example the therapeutic protein streptokinase

is analysed for the presence of potential MHC binding motifs and a method disclosed for the removal of a number of these from the molecule.

Streptokinase (SK) is a single chain protein of approximate molecular weight 47kDa that is produced by certain strains of β-haemolytic streptococci (Huang T.T. *et al.*, *Mol. Biol.* 2 197-205 (1989)). The protein has no inherent enzymatic activity but has considerable clinical importance owing to its ability to efficiently bind human plasminogen, potentiating its activation to plasmin and thereby promoting the dissolution of fibrin filaments in blood clots. Several studies have shown that SK is an effective thrombolytic agent in the treatment of coronary thrombosis, improving survival (ISIS-2 Collaborative Group, *Lancet* 2 349-360 (1988)) and preserving left ventricular function following myocardial infarction (ISAM Study Group, *N. Engl. J. Med.* 314 1465-1471 (1986); Kennedy J.W. *et al.*, *Circulation* 77 345-352 (1988)). Despite the undoubted therapeutic value of SK, the non-autologous origin of the protein is disadvantageous due to its immunogenicity in humans. The production of neutralising antibodies in the patient in generally limits the protein to a single use.

Method for identification of potential MHC class II binding motifs in streptokinase:

The sequence of streptokinase was identified from the GenBank database. The sequence with accession number S46536 was used throughout (Figure 5). The sequence was analysed for the presence of potential MHC class II binding motifs by computer aided comparison to a database of MHC-binding motifs as resident on world wide web site http://wehih.wehi.edu.au/mhcpep/.

Results of the "searching" process indicate the presence of 395 potential MHC class II binding motifs. Of these, 283 matched sequences identified in a database of human germline immunoglobulin variable region protein sequences. These epitopes were not considered further on the basis that immune responses in general are not mounted to

autologous circulating proteins such as immunoglobulins. This implies immunological tolerance to the potential T-cell epitopes present in the structure of the immunoglobulins (and indeed the majority of human proteins). Epitopes presented by non-autologous proteins such as SK which are identical or similar to motifs present in immunoglobulin proteins are likely also to be tolerated and in practice may be retained through the de-immunisation process.

Following subtraction of the human immunoglobulin protein germline motifs, the remaining 112 potential epitopes were analysed individually for similarity to non-immunoglobulin protein sequences. In practice, predicted anchor residues for each potential epitope were used in a consensus sequence search of human expressed proteins. The SwissProt and GenBank translated sequence databases were interrogated using commercially available software (DNAstar Madison, WI, USA). Epitopes identified in known circulating human proteins were not considered further and were therefore allowed to remain unchanged within the SK molecule. An example of one such rejected potential epitope is given by the sequence LLKAIQEQL at positions 79-87 in the SK protein. This sequence represents a predicted consensus binding motif for HLA-DR1*0101 with anchor residues underlined. Database searching using the consensus sequence LxxAxxxxL identifies >4000 entries in a human protein sub-set of the SwissProt database, including serum albumin protein (SwissProt accession number P02768). An example of an epitope where no match to a human protein considered to be in the general circulation was found is provided by sequence YVDVNTN at position 299-305 in the SK protein. This sequence represents a potential epitope for presentation by HLA-DR4*0401. Consensus sequence searching identifies < 50 human proteins containing this motif, of which many are intracellular proteins of differentiated tissues such as brain. These may be considered as not generally available to the immune system to gain tolerance and therefore identify this as a potential epitope for elimination according to the method of the present invention. Similarly, a further potential HLA-DR1*0101 binding motif was identified in the SK peptide sequence KADLLKAI at positions 76-83 of the SK

protein. This motif identifies < 150 human proteins in the same data set and was also identified for modification by the method of the present invention.

The net result of these processes was to identify those residues within the SK molecule which should be altered to eliminate potential MHC class II binding motifs. Individual amino acids within the predicted binding motifs were selected for alteration. With the object of maximising the likelihood of maintaining protein functional activity, in all cases conservative amino acid substitutions were chosen at any given site. A new (de-immunised) SK sequence was compiled (Figure 6) and further analysed by database comparison, as previously, for confirmation of successful elimination of potential MHC class II binding motifs.

Method for construction of de-immunised SK molecules:

PCR primers SK1 (5'-ggaattcatgattgctggacctgagtggctg) and SK2 (5'-tggatccttatttgtcgttagggtatc) were used to amplify the wild-type SK gene from a strain of *Streptococcus equisimililis* group C (ATCC accession number 9542). The resulting 1233bp fragment was cloned into pUC19 as a *Bam*HI-*Eco*RI restriction fragment using standard techniques (Sambrook J., Fritisch E.F. & Maniatis T. (eds.) in: Molecular Cloning a Laboratory Manual, Cold Spring Harbor Laboratory Press, NY, USA (1989)). The gene sequence was confirmed to be identical to database entries using commercially available reagent systems and instructions provided by the supplier (Amersham, Little Chalfont, UK). Site-directed mutagenesis was conducted using synthetic oligonucleotides and the "quick-change" procedure and reagents from Stratagene UK Ltd. Mutated (de-immunised) versions of the gene were confirmed by sequencing. Mutated SK genes were sub-cloned as *Eco*RI-*Bam*HI fragments into the bacterial expression vector pEKG-3 (Estrada M.P. *et al.*, *Bio/Technology* 10 1138-1142 (1992)) for expression of de-immunised SK. Recombinant protein was purified using a plasminogen affinity column according to the method of Rodriguez *et al.* (Rodriguez P. *et al.*, *Biotechniques* 7 638-641 (1992)). Fibrinolytic activity was assessed using the casein/plasminogen plate technique and the *in vitro* clot lysis assay

as described by Estrada *et al.* (Estrada *et al.*, *ibid.*).

Example 3

In the present example staphylokinase is analysed for the presence of potential MHC binding motifs and a method disclosed for the removal of a number of these from the molecule.

Staphylokinase protein from *Staphylococcus aureus* has recognised and well characterised profibrinolytic properties. Recombinant forms of the protein have been previously produced and *in vitro* and *in vivo* studies have indicated that the protein holds considerable promise for thrombolytic therapy (Sako, T., *Eur. J. Biochem.* 149 557-563 (1985); Schlott, B. *et al.*, *Biotechnology* 12 185-189 (1994); Collen, D. *et al.*, *Circulation* 94 197-462 (1996)). However, clinical use in humans has been limited due to the demonstrated immunogenicity of the protein in man (Collen, D. *et al.*, *Circulation* 95 463-472 (1997)). Availability of a non-immunogenic staphylokinase would have considerable importance as a potential agent for thrombolytic therapy.

The mature staphylokinase protein consists of a single polypeptide chain of 137 amino acids with approximate molecular weight 15.4kDa (Silence K. *et al.*, *J Biol. Chemistry*, 270 27192-27198 (1995)).

Method for identification of potential MHC class II binding motifs in staphylokinase:

The sequence of staphylokinase (sakSTAR) protein as given in table 1 of Collen *et al.* (Collen, D *et al.* (1996) *ibid.*) was used throughout. The protein sequences of the mature staphylokinase (Figure 7) were analysed for the presence of potential MHC class II binding motifs. Analysis was conducted by computer aided comparison to a database of MHC-binding motifs as resident on world wide web site http://wehih.wehi.edu.au/mhcpep/.

Results of the "searching" process indicate the presence of 128 potential MHC class II binding motifs. Of these, 91 matched sequences identified in a database of human germline immunoglobulin variable region protein sequences. These epitopes were not considered further. Epitopes presented by non-autologous proteins such as staphylokinase which are identical or similar to motifs present in immunoglobulin proteins are likely to be tolerated and in practice may be retained through the de-immunisation process.

Following subtraction of the human immunoglobulin protein germline motifs, the remaining 37 potential epitopes were analysed individually for similarity to non-immunoglobulin protein sequences. In practice, predicted anchor residues for each potential epitope were used in a consensus sequence search of human expressed proteins. The SwissProt and GenBank translated sequence databases were interrogated using commercially available software (DNAstar Madison, WI, USA). Epitopes identified in known circulating human proteins were not considered further and were therefore allowed to remain unchanged within the protein.

The net result of these processes was to identify those residues within the staphylokinase molecule which should be altered to eliminate potential MHC class II binding motifs. Individual amino acids within the predicted binding motifs were selected for alteration. With the object of maximising the likelihood of maintaining protein functional activity, in all cases conservative amino acid substitutions were chosen at any given site.

An altered staphylokinase sequence was compiled (Figure 8) and further analysed by database comparison, as previously, for confirmation of successful elimination of potential MHC class II binding motifs.

Method for construction of altered staphylokinase molecules:

A wild-type staphylokinase gene was synthesised under contract by Genosys

Biotechnologies Ltd (Cambridge, UK). The gene was constructed by PCR using long (80-mer) overlapping synthetic primers and the sequence as given by Collen, D *et al.* (Collen, D. *et al.*, (1996) *ibid.*) The synthetic gene was cloned as a 453 bp *Eco*RI-*Hin*dIII restriction fragment into bacterial expression vector pMEX (MoBiTec, Gottingen, Germany). The gene sequence was confirmed to be identical to database entries using commercially available reagent systems and instructions provided by the supplier (Amersham, Little Chalfont, UK). Altered (reduced immunogenicity) versions of the gene were engineered using site directed mutagenesis of the wild-type gene in pMEX. Short (18-mer) synthetic oligonucleotides and the "quick-change" procedure and reagents from Stratagene (Cambridge, UK) were used to create the variant genes. All variant gene sequences were confirmed by DNA sequencing.

Mutated staphylokinase genes were transformed into *E. coli* strain TG1 by standard techniques. A single transformed clone was selected and clones secreting active staphylokinase were selected using a fibrin plate assay (Astrup, T. *et al.*, *Arch. Biochem. Biophys.* 40: 346-351 (1952); Collen, D. *et al.*, *Fibrinolysis* 6: 203-213 (1992)). The best expressing clone was grown up and recombinant protein was purified from the culture supernatant using sequential column chromatography and methods as described previously (Collen, D. *et al.*, (1992) *ibid.*; Schlott *et al.*, *ibid.*).

Example 4

In the present example bryodin 1 is analysed for the presence of potential MHC binding motifs and a method disclosed for the removal of a number of these from the molecule.

The gene for bryodin 1 protein has recently been cloned from *Bryonia dionia*, a member of the *Cucurbitaceae* family of plants (Gawlak, S. *et al.*, *Biochemistry* 36 3095-3103 (1997)). Bryodin 1 is a type 1 ribosome inactivating protein. Studies using recombinant forms of the protein have indicated that bryodin 1 holds considerable promise for immunotoxin therapy for cancer and other diseases (Gawlak, S. *et al.*,

(1997) *ibid.*). However, clinical use in humans as with other immunotoxin agents, is likely to be curtailed due to immunogenicity of the protein in man. Availability of a non-immunogenic bryodin would have considerable importance as a potential component in immunotoxin based therapies.

The mature bryodin protein consists of a single polypeptide chain of 267 amino acids with approximate molecular weight 29kDa. The wild-type sequence is illustrated in Figure 9.

Method for identification of potential MHC class II binding motifs in bryodin 1: The sequence of bryodin 1 protein as given by Gawlak *et al.*, (Gawlak, S. *et al.*, (1997) *ibid.*) was used throughout. The protein sequence of the mature bryodin 1 (Figure 9) was analysed for the presence of potential MHC class II binding motifs. Analysis was conducted by computer aided comparison to a database of MHC-binding motifs as resident on world wide web site http://wehih.wehi.edu.au/mhcpep/.

Results of the "searching" process indicate the presence of 315 potential MHC class II binding motifs. Of these, 259 matched sequences identified in a database of human germline immunoglobulin variable region protein sequences. These epitopes were not considered further.

Following subtraction of the human immunoglobulin protein germline motifs, the remaining 56 potential epitopes were analysed individually for similarity to non-immunoglobulin protein sequences. The predicted anchor residues for each potential epitope were used in a consensus sequence search of human expressed proteins. The SwissProt and GenBank translated sequence databases were interrogated using commercially available software (DNAstar Madison, WI, USA). Epitopes identified in known circulating human proteins were not considered further and were therefore allowed to remain unchanged within the protein.

The net result of these processes was to identify those residues within the bryodin 1 molecule which should be altered to eliminate potential MHC class II binding motifs. For bryodin 1 protein, 13 potential epitopes were identified for elimination from the molecule. Individual amino acids within the predicted binding motifs were selected for alteration. With the object of maximising the likelihood of maintaining protein functional activity, in all cases conservative amino acid substitutions were chosen at any given site.

An altered bryodin 1 sequence was compiled (Figure 10) and further analysed by database comparison, as previously, for confirmation of successful elimination of potential MHC class II binding motifs.

Method for construction of altered bryodin 1 molecules:

A wild-type bryodin 1 gene was synthesised under contract by Genosys Biotechnologies Ltd (Cambridge, UK). The gene was constructed by PCR using long (80-mer) overlapping synthetic primers and the sequence as given by Gawlak *et al.*, (Gawlak, S. *et al.*, (1997) *ibid.*). The synthetic gene was cloned as a 843bp *Nco*I-*Eco*RI restriction fragment into a modified version of bacterial expression vector pET22b+ (Novagen, Madison, USA). The vector was modified to remove the pelB leader sequence which was previously shown to impede efficient expression of the bryodin 1 protein (Gawlak, S. *et al.*, (1997) *ibid.*). The modification was conducted by digestion with *Xba*I and *Nco*I to remove 107bp of DNA encompassing the pelB leader sequence. Elements of non-pelB sequence including the ribosome binding site sequence and the *Nco*I site, were restored in the vector by ligation of a linker molecule to the *Xba*I/*Nco*I free ends in the vector. The linker was formed by annealing complementary oligonucleotides:

L1f (5'-ctagaaataattttgtttaactttaagaaggagatatacatatgcc) and
L1r (5'-ccatggatatgtatatctccttcttaaagttaaacaaaattattt).

Oligonucleotides were supplied with phosphorylated ends by Genosys Biotechnologies Ltd (Cambridge, UK). Restriction digests, DNA purification, ligation reactions *etc.* were conducted using standard procedures and conditions recommended by the regent suppliers. Altered (reduced immunogenicity) versions of the gene were engineered using site directed mutagenesis of the wild-type gene in the modified pET22b vector. Short (18-mer) synthetic oligonucleotides and the "quick-change" procedure and reagents from Stratagene (Cambridge, UK) were used to create the variant genes. All variant gene sequences were confirmed by DNA sequencing.

Wild-type and mutated bryodin 1 gene variants were transformed into *E. coli* strain TG1 by standard techniques. A single transformed clone was selected for each gene and this clone used for sequence analysis. For expression work, bryodin 1 genes were transformed into *E. coli* strain BL21(λDE3) obtained from the ATCC. Recombinant bryodin 1 and variant bryodin 1 proteins were purified using methods described previously (Gawlak, S. *et al.*, (1997) *ibid.*). Following re-folding of the crude protein from inclusion bodies, purified bryodin 1 and variants were obtained > 95% pure using CM-Sepharose chromatography as previously (Gawlak, S. *et al.*, (1997) *ibid.*). Activity of the recombinant proteins was assessed using a cell-free protein synthesis inhibition assay and methods according to Siegall *et al.*, (Siegall, C.B. *et al.*, *Bioconjugate Chem.* 5 423-429 (1994)).

Example 5

In the present example the human interferon α2 protein is analysed for the presence of potential MHC binding motifs and a method disclosed for the removal of a number of these from the molecule.

Interferon alpha 2 (INA2) is an important glycoprotein cytokine expressed by activated macrophages. The protein has antiviral activity and stimulates the production of at least two enzymes; a protein kinase and an oligoadenylate synthetase,

on binding to the interferon alpha receptor in expressing cells. The mature INA2 protein is single polypeptide of 165 amino acids produced by post-translational processing of a 188 amino acid pre-cursor protein by cleavage of a 23 amino acid signal sequence from the amino terminus.

The protein has considerable clinical importance as a broad spectrum anti-viral, anti-proliferative and immunomodulating agent. Recombinant and other preparations of INA2 have been used therapeutically in a variety of cancer and viral indications in man (reviewed in Sen, G.G. and Lengyel P, *J. Biol. Chem.* 267 5017-5020 (1992)). However despite very significant therapeutic benefit to large numbers of patients, resistance to therapy in certain patients has been documented and one important mechanism of resistance has been shown to be the development of neutralising antibodies detectable in the serum of treated patients (Quesada, J.R. *et al., J. Clin. Oncology* 31522-1528 (1985); Stein R.G. *et al., New Eng. J. Med.* 318 1409-1413 (1988); Russo, D. *et al., Br. J. Haem.* 94 300-305 (1996); Brooks M.G. *et al., Gut* 30 1116-1122 (1989)). An immune response in these patients is mounted to the therapeutic interferon despite the fact that a molecule of at least identical primary structure is produced endogenously in man. In the present example, an interferon alpha 2 molecule with reduced potential immunogenicity in man is presented.

Method for identification of potential MHC class II binding motifs in human interferon alpha 2:

The sequence of INF2 was identified from the GenBank database. The sequence with accession number P01563 was used throughout. The protein sequences of the mature INF2 protein was identified and the sequence excluding the first 23 amino acid signal peptide analysed for the presence of potential MHC class II binding motifs (Figure 11). Analysis was conducted by computer using MPT ver1.0 software (Biovation, Aberdeen, UK). This software package conducts "peptide threading" according to the methods disclosed in WO-A-9859244. The software is able to provide an index of potential peptide binding to 18 different MHC class II DR alleles covering greater than

96% of the HLA-DR allotypes extant in the human population.

Results of the "peptide threading" process on the INF2, indicate the presence of a total of 18 individual potential epitopes. The epitopes map to five distinct clusters of overlapping epitopes encompassing residues 7-40 (cluster 1), residues 45-70 (cluster 2), residues 79-103 (cluster 3), residues 108-132 (cluster 4) and residues 140-163 (cluster 5). Each of the five clusters contain 5, 3, 4, 3 and 3 potential epitopes respectively.

In order to prioritise epitopes for removal, the epitope clusters were then mapped to the known structure-function features on the molecule. For INF2 evidence from homology modelling (Murgolo N.J. *et al.*, *Proteins: Structure, Function & Genetics* 17 62-74 (1993)), site directed mutagenesis (Tymms, M.J. *et al.*, *Antiviral Res.* 12 37-48 (1989); McInnes B. *et al.*, *J. Interferon Res.* 9 305-314 (1989)), cross-species chimaeric molecules (Ra, N.B.K. *et al.*, *J. Biol. Chem.* 263 8943-8952 (1988); Shafferman, A. *et al. J. Biol. Chem.* 262 6227-6237 (1987)), deletion mutants (Wetzel, R. *et al.*, In: *Interferons* New York Academic Press, pp819-823 (1982)) and serological mapping studies (Lydon N.B. *et al.*, *Biochemistry* 24 4131-4141 (1985); Trotta P.P. *et al.*,. In: *The Interferon System* Dianzani F & Rossi G.B (eds.) New York, Raven Press, 1985 pp231-235), suggest epitopes in clusters 1 and 2 are the highest priority targets for removal. With reference to the structural model of Murgolo et al (Murgolo N.J. et al (1993) *ibid.*), cluster 1 featuring 5 potential epitopes encompasses the functionally important helix A and the AB surface loop region of the INF2 molecule. This region is important for the antiviral activity and is involved in binding to the human INF2 receptor structure. Most significantly, epitopes for neutralising antibodies have been mapped to this region, specifically residues 10-11 of the helix A structural element (Lydon N.B. *et al.*, (1985) *ibid.*).

On this basis, cluster 1 epitopes at positions 7-19 and 13-25 were eliminated by substitution of threonine for leucine at position 15 ($L_{15}T$ using single letter codes).

This procedure was conducted interactively *in silico* using the MPTver1 package. Similarly, cluster 1 epitope at position 28-41 was eliminated by substitution $F_{27}Y$. This latter substitution had the concomitant effect of reducing the number of potential binding alleles for the overlapping epitope at position 22-34 from 13 to 11.

Cluster 2 epitopes extend from the AB loop through the helix B region and into the BC loop domain. Neutralising antibodies have been shown to bind in the BC loop (Lydon N.B. *et al.*, (1985) *ibid.*; Trotta P.P. *et al.*,. (1985) *ibid.*) therefore cluster 2 epitope at position 58-70 was also targeted for removal. Substitution $N_{65}L$ reduces potential immunogenicity in this region by eliminating 3 overlapping epitopes encompassing positions 61-77, which collectively are predicted to bind 5 different MHC DR alleles. Additionally, this substitution also reduces the number of different binding alleles at epitope 58-70 from 5 to 3.

Other epitope clusters (e.g. 3-5) map to either buried regions of the molecule or surface regions not involved in receptor binding and hence provide antigenic sites to which antibody responses are largely non neutralising.

Using the above method, an altered INF2 protein sequence was compiled containing 3 substitutions from the starting sequence and is depicted in Figure 12. This sequence is predicted to be significantly less immunogenic with respect to human MHC class II presentation. The areas of reduced immunogenicity focussed to regions of the molecule where antibody mediated neutralisation of the protein has been shown to occur with the potential to limit the clinical efficacy of the molecule as a therapeutic entity.

Method for construction of altered interferon alpha 2 molecule:

A wild-type INA2 gene was synthesised under contract by Genosys Biotechnologies Ltd (Cambridge, UK). The gene was constructed by PCR using long (80-mer)

overlapping synthetic DNA primers and the sequence given in GenBank accession number M29883. The synthetic gene was cloned as a 520bp *Eco*RI-*Hin*dIII restriction fragment into bacterial expression vector pMEX8 (MoBiTec, Gottingen, Germany). The gene sequence was confirmed to be identical to the desired gene using commercially available reagent systems and instructions provided by the supplier (Amersham, Little Chalfont, UK).

Altered (reduced immunogenicity) versions of the protein were constructed by site directed mutagenesis of the wild-type gene in pMEX8. Mutagenesis was conducted using short (18-mer) synthetic oligonucleotides obtained commercially (Genosys, Cambridge, UK) and the "quick-change" procedure and reagents from Stratagene (Cambridge, UK). Following site directed mutagenesis, DNA sequences of selected clones were confirmed as previously.

For expression of recombinant wild-type and recombinant variant INA2 proteins, pMEX8-INA2 expression plasmids were transformed into *E. coli* strain JA221 and cells grown and harvested using standard procedures (Sambrook J., Fritisch E.F. & Maniatis T. (1989) *ibid.*). Recombinant INA2 was prepared essentially as described previously (Grosfeld, H. *et al.*, in *Advances in Biotechnological Processes 4*, Mizrahi A. & Van Wezel A. L. eds., pp59-78 Alan R. Liss Inc, New York (1985)) with minor modifications. Briefly, following high speed centrifugation, the supernatant was concentrated by 60% saturation of ammonium sulphate and then chromatographed on a 2.7 x 68cm Sephadex G-75 column equilibrated with PBS plus 0.5M NaCl. Fractions (8ml) were collected at a 1min/ml flow rate. Pooled fractions were further purified using an immunoaffinity column as described previously (Grosfeld, H. *et al.*, (1985) *ibid.*). Purified proteins were assayed using SDS-PAGE analysis and functional activity (anti-viral activity) determined using biological assays as described previously (Shafferman, A. *et al.*, *J. Biol. Chem.* 262 6227-6237 (1987)).

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method of rendering an endogenous intracellular protein, or part of a protein, non-immunogenic, or less immunogenic, to a given species, the method comprising:

(a) determining at least part of the amino acid sequence of said endogenous intracellular protein;

(b) identifying in the amino acid sequence one or more potential epitopes for T cells ("T cell epitopes") which are found in said endogenous intracellular protein; and

(c) modifying the amino acid sequence to eliminate at least one of the T cell epitopes identified in step (b) thereby to reduce the immunogenicity of the protein or part thereof when exposed to the immune system of the given species.

2. A method as claimed in claim 1, wherein the given species is human.

3. A method as claimed in claim 1 or 2, wherein the T cell epitopes are identified by computation.

4. A method as claimed in any of the claims 1 to 3, wherein a T cell epitope is eliminated by alteration of one or more amino acids within the epitope itself.

5. A method as claimed in claim 4, wherein the alteration is an amino acid substitution.

6. A method as claimed in claim 5, wherein the amino acid substitution is made on the basis of similar size and/or charge.

7. A method as claimed in claim 4 or 5, wherein the amino acid substitution is made with reference to *in silico* protein modelling techniques.

8. A method substantially as hereinbefore described with reference to the examples and/or the preferred embodiments and excluding, if any, comparative examples.

Dated this tenth day of August 2004

Biovation Limited
Patent Attorneys for the Applicant:

F B RICE & CO

A
α-chain

ALYTLITPAVLRTDTEEQILVEAHGDSTPKQLDIFVHDFPRKQKTLFQTRVDMNPAGGMLVTP
TIEIPAKEVSTDSRQNQYVVVQVTGPQVRLEKVVLLSYQSSFLFIQTDKGIYTPGSPVLYRVF
SMDHNTSKMNKTVIVEFQTPEGILVSSNSVDLNFFWPYNLPDLVSLGTWRIVAKYEHSPENYT
AYFDVRKYVLPSFEVRLQPSEKFFYIDGNENFHVSITARYLYGEEVEGVAFVLFGVKIDDAKK
SIPDSLTRIPIIDGDGKATLKRDTFRSRFPNLNELVGHTLYASVTVMTESGSDMVVTEQSGIH
IVASPYQIHFTKTPKYFKPGMPYELTVYVTNPDGSPAAHVPVVSEAPHSMGTTLSDGTAKLIL
NIPLNAQSLPITVRTNHGDLPRERQATKSMTAIAYQTQGGSGNYLHVAITSTEIKPGDNLPVN
FNVKGNANSLKQIKYFTYLILNKGKIFKVGRQPRRDGQNLVTMNLHITPDLIPSFRFVAYYQV
GNNEIVADSVWVDVKDTCMGTLVVKGDNLIQMPGAAMKIKLEGDPGARVGLVAVDKAVYVLND
KYKISQAKIWDTIEKSDFGCTAGSGQNNLGVFEDAGLALTTSTNLNTKQRSAAKCPQPAN

B
β-chain

EIQMPTHKDLNLDITIELPDREVPIRYRINYENALLARTVETKLNQDITVTASGDGKATMTIL
TFYNAQLQEKANVCNKFHLNVSVENIHLNAMGAKGALMLKICTRYLGEVDSTMTIIDISMLTG
FLPDAEDLTRLSKGVDRYISRYEVDNNMAQKVAVIIYLNKVSHSEDECLHFKILKHPEVGFIQ
PGSVKVYSYYNLDEKCTKFYHPDKGTGLLNKICIGNVCRCAGETCSSLNHQERIDVPLQIEKA
CETNVDYVYKTKLLRIEEQDGNDIYVMDVLEVIKQGTDENPRAKTHQYISQRKCQEALNLKVN
DDYLIWGSRSDLLPTKDKISYIITKNTWIERWPHEDECQEEEFQKLCDDFAQFSYTLTEFGCP
T

C
γ-chain

DDNEDGFIADSDIISRSDFPKSWLWLTKDLTEEPNSQGISSKTMSFYLRDSITTWVVLAVSFT
PTKGICVAEPYEIRVMKVFFIDLQMPYSVVKNEQVEIRAILHNYVNEDIYVRVELLYNPAFCS
ASTKGQRYRQQFPIKALSSRAVPFVIVPLEQGLHDVEIKASVQEALWSDGVRKKLKVVPEGVQ
KSIVTIVKLDPRAKGVGGTQLEVIKARKLDDRVPDTEIETKIIIQGDPVAQIIENSIDGSKLN
SIPD

FIGURE 1

Protein sequence of mature cobra venom factor
A = alpha chain
B = gamma chain
C = beta chain

ALYTLITPAVLRTDTEEQILVEAHGDSTPKQLDIFVHDFPRKQKTLFQTRVDMNPAGGMLVTP
TIEIPAKEVSTDSRQNQYVVVQVTGTQVRLEKVVLLSYQSSFLFIQTDKGIYTPGSPVLYRVF
SMDHNTSKMNKTVIVEFQTPEGILVSSNSVDSLNFFWPYNLPDLVSLGTWRIVAKYEHSPENY
TAYFDVRKYVLPSFEVIVEPTEKFYYIDGNENFHVSITARYLYGEEVEGVAFVLFGVKIDDAK
ISLPESLKRIPIIDGDGKATLSRDTFRSRFPNLNELVGKSLYVSATVITESGSDMVVTEQSGI
HIVASPYQIHFTKTPKYFKPGMPYELTVFVTNPDGSPAAHVPVVVEAFHSMGTTLSDGTAKLI
LNIPLNAQSLPITVRTNHGDLPRERQATKSMTAIAYQTQGGSGNYLHVAITSTEIKPGDNLPV
NFNVKGNANSLAQIKYFTYLILNKGKIFKVGRQPREPGQDLVVLNLHITPDLIPSFRLVAYYT
LIGASGNNEIVADSVWVDVKDSCVGSIVVKGDNLIQMPGAAMKIKLEGDPGARVGLVAVDKAV
FVLNDKYKISQAKIWDTVVEKADIGCTAGSGQNNLGVFEDAGLALTTSTNLNTKQRSAAKCPQ
PAN

FIGURE 2

Protein sequence of an altered cobra venom factor α-chain.

DDNEDGFIADSDIISRSDFPKSWLWLTKDLTEEPNSQGISSKTMSFYLKDSITTWEVLAVSFT
PTKGICVADPFEVTVMKVFFIDLQMPYSVVKNEQVEIRAILHNYVNEDLYVRVELLYNPAFCS
ASTTGQRYRQQFPIKALSSRAVPFVIVPLEQGLHDVEVKAAVYHHFISDGVRKKLKVVPEGVQ
KSTVTIVKLDPRAKGVGGTQLEVIKARKLDDRVPDTEIETKIIIQGDPVAQIIENSIDGSKLN
SIPD

FIGURE 3

Protein sequence of an altered cobra venom factor γ-chain.

EIQMPTHKDLNLDITIELPDREVPIRYRINYENASLARTVETKLNQDFTVTASGDGKATMTIL
TFYNAQLQEKANVCNKFDLNVSVENIHLNAMGAKNTMILKICTRYLGEVDSTMTIIDISMLTG
FLPDAEDLTRLSKGVDRYISRYEVDNNMAQKVAVIIYLDKVSHSEDDCLHFKILKHFEVGFIQ
PGSVKVYSYYNLDESCTRFYHPDKGTGLLNKICIGNLCRCAGETCSSLNHQERVDVPLQIEKA
CETNVDYVYKTKLLRIEEQDGNDEYVMDVLEVIKQGTDENQRAKTHQYISQRKCQEALNLKVN
DHYLIWGSRSDLLPTKDKISYIIGKDTWVEHWPEEDECQEEEFQKLCDDFAQFSYTLTEFGCP
T

FIGURE 4

Protein sequence of an altered cobra venom factor β-chain.

IAGPEWLLDRPSVNNSQLVVSVAGTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKS
KPFATDSGAMPHKLEKADLLKAIQEQLIANVHSNDDYFEVIDFASDATITDRNGKVYFAD
KDGSVTLPTQPVQEFLLSGHVRVRPYKEKPIQNQAKSVDVEYTVQFTPLNPDDDFRPGLK
DTKLLKTLAIGDTITSQELLAQAQSILNKTHPGYTIYERDSSIVTHDNDIFRTILPMDQE
FTYHVKNREQAYEINKKSGLNEEINNTDLISEKYYVLKKGEKPYDPFDRSHLKLFTIKYV
DVNTNELLKSEQLLTASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDD
TNRIITVYMGKRPEGENASYHLAYDKDRYTEEEREVYSYLRYTGTPIPDNPNDK

FIGURE 5

Protein sequence of streptokinase from *Streptococcus equisimilis*

IAGPEWLLDRPSVNNSQLVVSVAGTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKS
KPFATDSGAMPHKLEKADLLKAKQEQLIANVHSNDDYFEVIDFASDATITDRNGKVYFAD
KDGSVTLPTQPVQEFLLSGHVRVRPYKEKPIQNQAKSVDVEYTVQFTPLNPDDDFRPGLK
DTKLLKTLAIGDTITSQELLAQAQSILNKTHPGYTIYERDSSIVTHDNDIFRTILPMDQE
FTYHVKNREQAYEINKKSGLNEEINNTDLISEKYYVLKKGEKPYDPFDRSHLKLFTIKFV
DVNTNELLKSEQLLTASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDD
TNRIITVYMGKRPEGENASYHLAYDKDRYTEEEREVYSYLRYTGTPIPDNPNDK

FIGURE 6

Protein sequence of an altered streptokinase molecule.

SSSFDKGKYKKGDDASYFEPTGPYLMVNVTGVDSKGNELLSPHAVEFPIKPGTTLTKEKIEYY
VEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSFPITEKGFVVPDLSEHIKNPGFN
LITKVVIEKK

FIGURE 7

Protein sequence of mature staphylokinase.

SSSFDKGKYKEGDDASQFEPTGPYLMVNVTGVDSAGNALLSPHYVEFPIKPTTLTEERIKYYV
EWALDATAYAAFAVVELDPSARVEVTYYDKNKKKEETKSFPITEKGFVVPDTSEHIKNPGFNL
FTKVVIEKK

FIGURE 8

Protein sequence of an altered staphylokinase molecule.

DVSFRLSGATTTSYGVFIKNLREALPYERKVYNIPLLRSSISGSGRYTLLHLTNYADETISVA
VDVTNVYIMGYLAGDVSYFFNEASATEAAKFVFKDAKKKVTLPYSGNYERLQTAAGKIRENIP
LGLPALDSAITTLYYYTASSAASALLVLIQSTAESARYKFIEQQIGKRVDKTFLPSLATISLE
NNWSALSKQIQIASTNNGQFESPVVLIDGNNQRVSITNASARVVTSNIALLLNRNNIAAIGED
ISMTLIGFEHGLYGI

FIGURE 9
Protein sequence of wild-type bryodin 1.

DVSFSMSGATTTSYGVFVKNLREALPFERKVYNIPLLRSSISGSGRYTLLHLTNYADETISVA
VDVTNVYIMGYLAGDVSYFFNEASATEAAKFVFKDAKKKVTLPYSGNHERLQTAAGKIRENIP
LGLPALDSAITTLYYYTASSAASALLVLIQSTAESARFKFIEQQIGKRVDKTFLPSLATISLE
NNWSALSKQIQIASTNNGQFESPVVLVDGNNQSVSITNASARVVTSNVALLLNRNNIAAVGED
ISMTLIGFEHGLYGI

FIGURE 10
Protein sequence of an altered bryodin 1 molecule.

CDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMIQQI
FNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVRKYFQRI
TLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE

FIGURE 11
Protein sequence of mature human interferon alpha 2

CDLPQTHSLGSRRTTMLLAQMRKISLFYCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMIQQI
FLLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVRKYFQRI
TLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE

FIGURE 12
Protein sequence of an altered human interferon alpha 2 molecule